Effect of competing mortality risks on predictive performance of the QRISK3 cardiovascular risk prediction tool in older people and those with comorbidity

Impact of competing mortality risks on predictive performance of the QRISK3 cardiovascular risk prediction tool in older people and those with comorbidity: external validation population cohort study

Dr Shona Livingstone, Division of Population Health and Genomics, University of Dundee

Dr Daniel R Morales, Division of Population Health and Genomics, University of Dundee

Prof Peter T Donnan, Division of Population Health and Genomics, University of Dundee

Prof Katherine Payne, Division of Population Health, Health Services Research & Primary Care

Dr Alexander Thompson, Division of Population Health, Health Services Research & Primary Care

Dr Ji-Hee Youn, Division of Population Health, Health Services Research & Primary Care

Prof Bruce Guthrie, Usher Institute, University of Edinburgh

Corresponding author:
Bruce Guthrie, Doorway 3, Old Medical School, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG
bruce.guthrie@ed.ac.uk

Keywords: Cardiovascular risk; primary prevention; risk-prediction; external validation; QRISK3; competing mortality risk

Word count: 3422

Abstract

Background: Cardiovascular disease (CVD) primary prevention is guided by risk-prediction tools, but these rarely account for the risk of dying of other conditions (competing mortality risk).

Methods: Retrospective population cohort study to externally validate the UK-recommended QRISK3 CVD risk-prediction tool in Clinical Practice Research Datalink data. QRISK3 predicted 10-year CVD risk was compared to observed 10-year risk in the whole population and important subgroups. QRISK3 discrimination and calibration was examined with and without account for competing risks.

Findings: 1,484,597 women with 42,451 incident CVD events (4.9/1000 person-years follow-up), and 1,420,176 men with 53,066 incident CVD events (6.7/1000) were included, with median follow-up of 5.0 years. Non-CVD death rose markedly with age (eg 0.4% of women and 0.5% of men aged 25-44 had non-CVD death vs 20.1% and 19.6% aged 75-84). QRISK3 discrimination in the whole population was excellent (Harrell's C=0.865 in women, 0.834 in men) but was poor in important subgroups (eg C<0.65 in all subgroups aged 65+). Ignoring competing risks, QRISK3 calibration in the whole population and in younger people was excellent, but there was significant over-prediction in older people. Accounting for competing risks, QRISK3 systematically over-predicted CVD risk, particularly in older people, and in people with high comorbidity.

Interpretation: QRISK3 performs well at whole population level when ignoring competing mortality risk. It performs considerably less well in important subgroups including older people and people with comorbidity, and less well again after accounting for competing mortality risk.

Funding: National Institute for Health Research HS&DR Programme 15/12/22

Research in Context

Evidence before this study

Guidelines for the primary prevention of cardiovascular disease (CVD) usually recommend risk-stratified treatment, where decisions to start long-term medication to prevent future CVD events are guided by estimation of CVD risk with treatment offered if patients exceed a particular risk threshold. Recommended risk-prediction tools vary by country, reflecting differences in CVD risk factors and incidence. The recommended risk-prediction tool in the UK is QRISK, but two criticisms of recommended tools are that (1) They often do not predict risk well in older people and people with multimorbidity; and (2) They do not account for competing mortality risk (the risk of dying from non-CVD causes). We searched Pubmed from inception to 8/1/21 for observational studies examining competing mortality risks in people with CVD or in the context of incident CVD risk-prediction. Estimated over-estimation of rates of CVD during follow-up has been found in the context of incident (first ever) CVD in the whole general population and high-risk populations such as people with atrial fibrillation, and of further CVD-related events in people with established CVD. The degree of over-estimation of CVD varies with the population, and is believed to be higher in older people where competing mortality risk is higher but is not usually taken into account by CVD-risk prediction tools.

Added value of this study

This external validation of the QRISK3 CVD risk-prediction tool found excellent discrimination and very good calibration at whole general population level. Discrimination and calibration were poor to moderate in older people and in people with high numbers of co-existing long-term conditions ('high comorbidity'), particularly after accounting for competing mortality risks and there was systematic over-prediction in older people and in people with high comorbidity.

Implications of all the available evidence

CVD risk-prediction models need to be validated in older people and in people with high comorbidity. Better CVD risk-prediction models are needed to stratify people potentially eligible for primary preventive treatments. Clinicians should consider competing mortality risks and non-CVD life expectancy when discussing statin initiation for primary prevention with older people and people with high comorbidity.

Introduction

Cardiovascular disease (CVD) remains a major cause of morbidity and mortality worldwide despite falling incidence in most developed countries. Guidelines for the primary prevention of CVD usually recommend the use of risk-prediction tools to target treatment for people above a specified threshold of predicted risk. There has been a progressive reduction in the risk threshold recommended in relation to statin prescription for primary prevention. In England and Wales, the 10-year CVD risk threshold at which treatment is recommended changed from 20% to 10% in 2014,[1] compared to a 7.5% threshold in current US guidelines.[2] These reductions reflect both increasing evidence of statin effectiveness for primary prevention, and falling prices making statins more cost-effective at lower levels of baseline risk (and therefore lower levels of expected absolute benefit). However, age is the most important predictor of CVD risk, so most people will exceed these lower thresholds at some point in early old age irrespective of other risk factors. ~~The implications of effectively recommending medication for all older people has proved controversial.[3]~~

Risk-stratified guideline recommendations rely on being able to accurately predict the risk of CVD events. Risk-prediction tools recommended in different countries vary, reflecting variation in CVD risk factors and incidence in different countries. In England and Wales, NICE recommend the QRISK2 risk-prediction tool,[1] which has been internally and externally validated in UK primary-care datasets, and found to have excellent discrimination and calibration at whole population level.[3~~-5~~] A new version of the QRISK tool (QRISK3) has been derived and internally validated using the same methodology in the QRESEARCH database. The main changes in QRISK3 were to add a number of new conditions to the prediction equation.[4] In QRISK3 internal validation in the original dataset, overall model discrimination was excellent, and calibration excellent in younger people and very good in older people,[4] but external validation is required before recommending any prediction tool for routine use.[5,6,7]

However, there are additional concerns about risk-prediction that are not directly addressed by conventional external validation. In particular, people who are very likely to die from non-CVD conditions may have little potential benefit from statins but at least some risk of harm from treatment. The issue is one of competing risk, which in this context arises when an individual is at risk of dying from conditions other than CVD. These are obvious at the extreme - taking a statin is clearly futile in someone at the very end of life. However, even smaller levels of competing risk can lead to systematic over-prediction of CVD risk in people at higher risk of dying from another cause, including older people and those with multimorbidity.[8,9] This is because survival analyses where data are censored usually assume that those lost to follow-up have the same risk of the outcome as those who remain in follow-up (for example, if using the Kaplan-Meier estimator). This assumption is

clearly incorrect if someone dies of another condition (competing mortality) since a dead person cannot experience a CVD event.[10] The aim of this analysis was to externally validate QRISK3 and examine the effect of competing risk on predictive performance.

Methods

Data source and population. We externally validated QRISK3 in Clinical Practice Research Datalink (CPRD) Gold),[11,12] which is non-overlapping with the derivation dataset, although similar in its inclusion of linked primary care, hospital and mortality data. To be included, patients had to: be permanently registered with a general practice contributing up-to-standard data for at least one year and with linkage to Hospital Episode Statistics discharge and Office of National Statistics mortality data; be aged ≥25 years and <85 years with no prior history of CVD; and have no history of prior statin treatment. Cohort entry was the latest of these dates on or after 01/01/04. Cohort exit was the date of the earliest of: first CVD event; death; prescription of a statin; deregistration from the general practice; date of the last data collection from the practice; or the end of the study on 31/3/16. All outcomes and predictors are as recorded during routine clinical care, and are therefore recorded blind to the study hypothesis. The study size is determined by the data available in CPRD which was considered sufficient,[5] and no formal power calculation was done.[5]

Outcomes. A first CVD event was defined as the earliest recording of any fatal or non-fatal coronary heart disease (CHD), ischaemic stroke, or transient ischaemic attack. Fatal CVD events were identified from ICD-10 codes recorded in Office of National Statistics death registration. Non-fatal events were identified either in GP EHRs (using Read codes, the standard coding system used in UK general practice) or Hospital Episode Statistics (HES) discharge diagnoses (ICD-10 codes). Read and ICD-10 codes defining outcomes are those used in QRISK3 derivation,[4] and are listed in supplementary table S1

Prediction model. We implemented the published QRISK3-2017 prediction model (under GNU Lesser General Public Licence v3) with some exceptions, namely: (1) We chose a later cohort entry date (1/1/04 rather than 1/1/98); (2) We handled cholesterol missingness differently (if no values were available at baseline, QRISK3 derivation allowed cholesterol values from *after* the index date to be used if they were before any event; we only included values recorded before the index date to avoid using future information in prediction); and (3) We evaluated the Townsend deprivation score as the median of the vigintile (equal 20th) of score that an individual lived in, as individual values were not available. Predictor code sets used and methods of data handling are detailed in supplementary tables S2-S4.

Comorbidity. For each patient at baseline, we additionally calculated a modified Charlson Comorbidity Index (mCCI) based on primary care Read codes (modified in that CVD could not contribute to the score as all participants are CVD-free at baseline).[13] mCCI was not used in prediction, but was used to stratify the population to examine discrimination and calibration by mCCI score (grouped into 0, 1, 2, and 3+).

Missing data. Supplementary table S5 details how missing data were treated for all variables, and the extent of missingness. As with QRISK3 derivation, patients with missing Townsend deprivation score were excluded from the cohort, those with missing ethnicity were assumed to be white, and multiple imputation was used for missing body mass index (BMI), total cholesterol:HDL cholesterol ratio (TC:HDL), systolic blood pressure (SBP), SBP variability, and smoking status. For the variables listed in supplementary table S5, Multivariate Imputation by Chained Equations[14] was used to generate five imputed datasets. Analyses of these datasets were combined using Rubin's rules[15] to give summary point estimates with confidence limits that reflect the added uncertainty associated with imputing missing values.

Statistical methods. The 10-year risk of suffering a cardiovascular event was calculated for each patient using the published QRISK3 equation without recalibration. The performance of the risk score was assessed by examining discrimination and calibration.

Discrimination is the ability of the risk score to differentiate between patients who experience the event of interest during the study and those who do not. We used the truncated version of Harrell's C-statistic to only include pairs where the earliest survival time is no later than 10 years after entry. Where there is considerable censoring, Harrell's C-statistic may be optimistic. We therefore carried out a sensitivity analysis using a weighted C-statistic accounting for probability of censoring.[16] A C-statistic of 0.5 indicates discrimination that is no better than chance, whereas a C-statistic of 1 indicates perfect discrimination. Two additional measures of discrimination were calculated, the D statistic of Royston and Sauerbrei (based on the separation in event free survival between those patients with predicted risk scores above and below the median; higher values indicate greater discrimination),[17] and a related R-squared statistic estimating explained variation in the context of censored survival data.[18]

Calibration refers to how closely the predicted and observed probabilities agree at group level. This was assessed for equally-sized groups of participants ranked by predicted risk. Calibration of the risk score predictions was assessed by plotting observed proportions versus predicted probabilities. Plots were generated separately by sex, for all patients and for pre-specified subgroups of age and mCCI based on summary statistics pooled across the imputed datasets.

The following summary statistics and their standard errors were obtained by decile of predicted risk score and for each imputed dataset in turn: non-parametric measures of observed risk or proportions of patients with a CVD event, the Kaplan-Meier estimator (the conventional measure ignoring competing risks) and the Aalen Johansen estimator (an extension to allow for competing events, non-CVD death in this case),[19] and the mean predicted risk score. All models were fitted in R-4.0.0 and STATA 11.2.

Findings

There were 1,648,746 women aged 25-84 with linkage to HES and ONS, of whom 164,129 (10.0%) were excluded because of prior CVD (4.7%), prior statin prescribing (5.1%), or missing deprivation score (0.2%). There were 1,621,535 men aged 25-84 with linkage to HES and ONS, of whom 201,359 (12.4%) were excluded because of prior CVD (6.9%), prior statin prescribing (5.3%) or missing deprivation score (0.2%). 1,484,597 women and 1,420,176 men were therefore included in the study.

Baseline characteristics of the study population are shown in table 1, compared to the QRISK3 internal validation cohort.[4] Across most characteristics, the two cohorts were similar, but in this study the prevalence of treated hypertension and current smoking was higher, and recording of a family history of CHD was lower. Supplementary table S5 details missing data, which was less frequent in this study for ethnicity, similar for systolic blood pressure and body mass index, and more frequent for total cholesterol:HDL cholesterol ratio, systolic blood pressure variability, and smoking status.

There were 42,451 incident cases of CVD observed in women in 8,594,620 years of follow-up (4.9 [95%CI 4.89-4.99] per 1000 person-years), compared to 53,066 incident cases in men in 7,896,704 years of follow-up (6.7 [95%CI 6.66-6.78] per 1000 person-years). CVD incidence rose progressively with age (supplementary table S6), and was moderately lower than that observed in QRISK3 derivation[4].

Follow up time and censoring. Median follow up in the whole cohort was 5.0 years, with 22.1% of patients still in the cohort and CVD event-free at 10 years. Table 2 shows the status of all patients at 10 years. In women, CVD events occurred in 2.6% by 10 years and non-CVD deaths in 2.8%, compared to 3.5% and 2.7% respectively in men. Censoring due to statin initiation was more common (8.6% of women, 10.2% of men), but almost two-thirds of both men and women were censored due to deregistration or by <10 years follow-up before the end of the study on 31/3/16. Patterns of censoring were markedly different by age and by modified Charlson Comorbidity Index (mCCI). Censoring due to statin initiation rapidly increased with age, peaking in 65-74 year olds

where approximately a quarter of men and women started statins for primary prevention during follow-up. Censoring due to deregistration or having less than 10-years follow-up by the end of study was markedly more common in younger people, whereas censoring due to non-CVD death was markedly more common in older people. Similar patterns to age were seen with mCCI

Discrimination. Overall discrimination was excellent and very similar to QRISK3 internal validation[4] (for women, Harrell's C-statistic=0.865 external validation vs 0.880 QRISK3 internal validation, D=2.43 vs 2.49, R^2 58.5% vs 59.6%; for men Harrell's-C 0.834 vs 0.858, D=2.10 vs 2.26, R^2=51.3% vs 55.0%). However, discrimination varied markedly within the age-group and mCCI categories, with discrimination being best in the youngest (25-44 years) and least multimorbid (mCCI=0) groups, and worst in the oldest (75-84 years) and most multimorbid (mCCI=3+) groups. For example, in women aged 75-84, Harrell's C statistic=0.611, and for men aged 75-84, Harrell's C-statistic=0.585, representing poor discrimination, with only moderate discrimination observed in women and men aged 65-74. Sensitivity analysis using a censoring-adjusted C-statistic found somewhat lower discrimination, but did not alter the overall interpretation (supplementary table S7)

Calibration in women. Figure 1 and Supplementary Figures S1-S2 show calibration plots for women. Ignoring competing mortality risks (left-hand plots), calibration is excellent for all women, and for women aged 25-44 years. However, QRISK3 progressively over-predicted CVD risk in older age-groups. Stratified by mCCI, there is evidence of some over-prediction in the least comorbid (mCCI=0), but of poor calibration and under-prediction in those with high comorbidity (mCCI≥3).

When competing mortality risks are accounted for (right-hand plots), there is over-prediction of risk at higher levels of predicted CVD risk in all women. The same pattern of increasing over-prediction with increasing age is observed, but with greater magnitude and calibration is poor in older age groups. Although there is some observed under-prediction of risk in those with mCCI≥3 for those at lower predicted CVD risk, the overall pattern is of over-prediction of CVD risk which increases with comorbidity, and calibration is poor in the most co-morbid (mCCI=2 and mCCI≥3).

Calibration in men. Figure 2 and Supplementary Figures S3-S4 show calibration plots for men. Ignoring competing mortality risks (left-hand plots), calibration is excellent for all men although with somewhat greater over-prediction at higher levels of predicted CVD risk than in women. Calibration is excellent for men aged 25-44 years, but QRISK3 progressively over-predicted CVD risk in older age-groups. Stratified by mCCI, there is evidence of some over-prediction in the least comorbid (mCCI=0), but of poor calibration and under-prediction in those with high comorbidity (mCCI≥3).

When competing mortality risks are accounted for (right-hand plots), there is over-prediction of risk at higher levels of predicted CVD risk in the whole population. Calibration is poor with large over-prediction in older age-groups. Although there is some observed under-prediction of risk in those with mCCI≥3 for those at lower predicted CVD risk, the overall pattern is of over-prediction of CVD risk which increases with comorbidity, and calibration is poor in the most co-morbid (mCCI=2 and mCCI≥3).

Discussion

This external validation study finds that at whole population level, QRISK3 has excellent discrimination (the ability of the model to distinguish people at higher or lower risk; women c=0.86, men c=0.83), but discrimination was only poor to moderate in people aged 75-84 (women c=0.61, men c=0.58) and moderate to good in people with high levels of comorbidity (mCCI≥3; women c=0.74, men c=0.70). Calibration (the extent to which predicted and observed event rates are similar) was excellent in the whole population when ignoring competing mortality risks but there was evidence of systematic under-prediction after competing risks were accounted for. Calibration was considerably worse in older people and in those with higher levels of comorbidity where QRISK3 systematically over-predicted risk, particularly after competing mortality risks were accounted for.

At whole population level, the QRISK3 CVD risk-prediction model does appropriately sort the whole population into groups with varying levels of cardiovascular risk (with some small over-prediction), but the model performs relatively poorly in older people and people with high comorbidity, in part because of high competing mortality risk.

Strengths of the study include conduct consistent with methodology recommendations,[6,20] comprehensive detailing of all code sets used to facilitate replication, and explicit consideration of prediction in subgroups and competing mortality risks.

Limitations largely reflect problems common to all studies using routine GP data, including the original QRISK3 derivation.[21] The prevalence of missing data for key predictors is high. Like the original QRISK3 derivation, we used multiple imputation for missing data, but in this context the assumption that data are missing at random is a strong one, but balanced against the use of representative population data. All recent QRISK models have also used 1/1/98 as the index date (the earliest that patients can enter the study). Much observed follow-up in model derivation is therefore historical,[21] and there is a trade-off between using an index date in the distant past (when CVD incidence is higher than now) or a more recent index date (in which case more patients are excluded because of prior statin use). Our choice of a more recent index date may partly explain why QRISK3 is observed to over-predict in our validation. Deriving clinical prediction tools on increasingly

historical data is likely biased,[21] but using more recent data with greater rates of prior statin initiation may also be biased. There is no clearly optimal resolution to this dilemma. Finally, loss to follow-up before a CVD event was common which is relevant to model assumptions about censored patients. We specifically examined the impact of censoring due to non-CVD death, but it is also an assumption that those who deregister from practices have the same CVD risk as those who do not. That seems likely to be true for younger people, but less so for older people where change of address will be more commonly driven by change in health status (eg moving to sheltered housing or a care home).

Previous external validations of previous QRISK tools have also found excellent discrimination and calibration at the whole population level when competing mortality risks were ignored (ie answering the question 'what is the risk of CVD assuming that this person does not die of anything else in the next 10 years?').[3] QRISK3 therefore 'works' at whole population level when considered in its own terms (ignoring competing risks), but even so, discrimination and calibration were poor in those aged 75-84 years, and only moderate in those aged 65-74 years and in those with the highest levels of comorbidity (mCCI=3).

There was greater over-prediction at whole population level once competing mortality risk was accounted for (ie answering the question 'what is the risk of CVD allowing for the risk of this person dying of something else first?'). Calibration was notably poorer once competing risk was accounted for, particularly in older and more comorbid patients. These findings are consistent with other studies examining the impact of competing risks on estimated CVD risk in people without CVD,[9,10,22,23] with established CVD,[24] and in other contexts including stroke risk in people with atrial fibrillation.[25,26] QRISK2 has also been shown to systematically over-predict CVD risk in a contemporary population of people with type 2 diabetes, with increasingly poor discrimination with increasing age, highlighting that good performance at whole population level does not necessarily mean good performance in important subgroups.[27]

At population level, QRISK3 does segment the population into groups which largely have the risk of CVD predicted (supporting its use to guide risk-stratified treatment decisions). However, this overall assessment of prediction performance is largely driven by good performance in younger people with fewer co-existing long-term conditions. For older people and people with more long-term conditions, prediction is poor to fair, particularly when competing risks are accounted for. The lower levels of over-prediction in other age and comorbidity groups will also sometimes change treatment recommendations in younger and less comorbid people. Similar issues likely apply to other CVD risk-prediction models which do not account for competing risk. We believe that predicting CVD events

without accounting for risk of death from other causes is misleading in people at high risk of non-CVD death. Clinicians should therefore carefully consider life expectancy related to other conditions when discussing long-term cardiovascular primary preventive treatment.

There are several areas where further research would be beneficial. CVD causes a large proportion of deaths, which will reduce the impact of competing risks. There is a need for further studies examining the impact of competing risk when predicting less commonly fatal conditions, where the impact on predictive performance is likely to be greater. It is also uncertain whether a better approach to CVD prediction would be to create separate models for important subgroups of age and comorbidity (as is already done for men and women). The relative merits of omnibus vs smaller subgroup models needs research, as does better quantifying the uncertainty at individual level of risk-prediction tools which perform well at population level.[28] A weakness of existing UK primary care datasets in deriving risk-prediction rules is the large loss to follow-up where there is a long time-horizon for risk-prediction. This study has examined the impact of competing risk, but other loss to follow-up due to practice deregistration is likely to create over-prediction in at least some population subsets. External validation in large geographical populations with less loss to follow-up (such as SAIL Databank in Wales) would be valuable, as would larger-scale data federation to derive and validate new risk-prediction tools for comparison with QRISK3 and other prediction models.[21]

In conclusion, QRISK3 performs well at whole population level, but systematically over-predicts CVD risk in older people and people with high comorbidity. Clinicians should consider broader impacts on life expectancy when discussing statin initiation for primary prevention in older people and people with high comorbidity in whom CVD risk is likely over-predicted. Better calibrated prediction models are needed in these groups.

Author contribution

The study was conceived of and designed by BG, KP, DRM, PTD, KP and AT who obtained the funding. All authors contributed to study design and interpretation. SL, BG, DRM and PTD led data management and SL led analysis supported by BG, DRM and PTD. SL and BG drafted the paper, which all authors reviewed and edited. SL, BG and DRM verified the underlying data.

Competing interests

None to declare

Ethics approval

The study was approved by the Clinical Practice Research Datalink Independent Scientific Advisory Committee protocol 16_248

Funding

This study/project is funded by the National Institute for Health Research (NIHR) Health Services and Delivery Research Programme (project reference 15/12/22). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The authors had full and sole access the data, and the funder had no role in the conduct of the research or the decision to publish.

Data sharing

The data controller is the Clinical Practice Research Datalink (CPRD), and under the data licence granted, the authors are not allowed to share data. Researchers can apply to CPRD directly for access to the raw data.

Table 1: Baseline data in external validation cohort and in original QRISK3 internal validation cohort[4]

	Women external validation cohort N=1484597	Women original QRISK3 internal validation cohort N=1360457	Men external validation cohort N=1420176	Men original QRISK3 internal validation N=1310841
Mean (SD) Age (years)	46.0 (15.3)	43.3 (15.3)	44.8 (13.9)	42.6 (13.8)
Mean (SD) Body mass index	25.9 (5.7)	25.4 (5.1)	26.6 (4.7)	25.9 (4.2)
Mean (SD) Total cholesterol:HDL cholesterol ratio	3.7 (1.1)	3.6 (1.2)	4.4 (1.3)	4.4 (1.3)
Mean (SD) Systolic blood pressure (mmHg)	125.4 (18.0)	123.1 (18.1)	131.1 (16.2)	128.8 (16.2)
Mean (SD) Systolic blood pressure variability	10.0 (5.7)	9.3 (6.1)	10.3 (6.2)	9.9 (6.8)
Ethnicity No. (%)				
White or not recorded	1363146 (91.8)	1218391 (89.6)	1336,221 (94.1)	1171281 (89.4)
Indian	22488 (1.5)	23146 (1.7)	15322 (1.1)	26479 (2.0)
Pakistani	9550 (0.6)	10919 (0.8)	6647 (0.5)	14787 (1.1)
Bangladeshi	2594 (0.2)	8738 (0.6)	2145 (0.2)	11914 (0.9)
Other Asian	13697 0.9)	17078 (1.3)	9973 (0.7)	15966 (1.2)
Black Caribbean	9505 (0.6)	13142 (1.0)	6687 (0.5)	10642 (0.8)
Black African	18804 (1.3)	27678 (2.0)	12822 (0.9)	25251 (1.9)
Chinese	6739 (0.5)	8992 (0.7)	3503 (0.2)	6098 (0.5)
Other	38074 (2.6)	32373 (2.4)	26829 (1.9)	28423 (2.2)
Smoking status No. (% of non-missing)				
Non-smoker	707774 (59.8)	706671 (51.9)	478671 (49.0)	512252 (39.1)
Former smoker	217404 (18.4)	194545 (14.3)	216883 (22.2)	196459 (15.0)
Light smoker	85277 (7.2)	154565 (11.4)	75260 (7.7)	177693 (13.6)
Moderate smoker	111690 (9.4)	74933 (5.5)	112411 (11.5)	84914 (6.5)
Heavy smoker	62236 (5.3)	38218 (2.8)	93457 (9.6)	64107 (4.9)
FH of CHD in first degree relative <60 years	97624 (6.6)	164023 (12.1)	75237 (5.3)	123039 (9.4)
Type 1 diabetes	3752 (0.3)	3351 (0.2)	4843 (0.3)	3932 (0.3)
Type 2 diabetes	17022 (1.1)	15872 (1.2)	21077(1.5)	19318 (1.5)
Treated hypertension	115944 (7.8)	77694 (5.7)	82768 (5.8)	56920 (4.3)

Rheumatoid arthritis	12702 (0.9)	15139 (1.1)	4724 (0.3)	7055 (0.5)
Atrial fibrillation	8199 (0.6)	5229 (0.4)	10620 (0.7)	6874 (0.5)
Chronic kidney disease (stage 3, 4 or 5)	6918 (0.5)	6949 (0.5)	5659 (0.4)	4232 (0.3)
Migraine	117692 (7.9)	89504 (6.6)	41471 (2.9)	36141 (2.8)
Corticosteroid use	20674 (1.4)	31775 (2.3)	11824 (0.8)	18634 (1.4)
HIV/AIDS	289 (0.02)	1595 (0.1)	445 (0.03)	2945 (0.2)
Systemic lupus erythematosus	1725 (0.1)	1349 (0.1)	165 (0.01)	134 (0.0)
Atypical antipsychotic use	8469 (0.6)	6268 (0.5)	8336 (0.6)	6597 (0.5)
Severe mental illness	110799 (7.5)	94724 (7.0)	57264 (4.0)	57830 (4.4)
Erectile dysfunction diagnosis or treatment	N/A	N/A	39,264 (2.8)	31136 (2.4)

Table 2: Follow-up and censoring events at ten years

	No entering study cohort	Non-fatal CVD No. (%)	CVD death No. (%)	Censored because non-CVD death No. (%)	Censored because started statin No. (%)	Censored because deregistered or end of study (31/3/16) before ten years FU No. (%)	At least ten years follow-up No. (%)
All women	1484597	34047 (2.3)	5001 (0.3)	40839 (2.8)	128183 (8.6)	926832 (62.4)	349695 (23.6)
All men	1420176	42675 (3.0)	6471 (0.5)	38226 (2.7)	145482 (10.2)	895421 (63.1)	291901 (20.6)
Women							
Aged 25-44	813157	3064 (0.4)	124 (0.0)	3250 (0.4)	14076 (1.7)	612336 (75.3)	180307 (22.2)
Aged 45-64	465484	10825 (2.3)	671 (0.1)	11101 (2.4)	68552 (14.7)	242367 (52.1)	131968 (28.4)
Aged 65-74	121267	8958 (7.4)	1142 (0.9)	9454 (7.8)	32139 (26.5)	43549 (35.8)	26205 (21.6)
Aged 75-84	84689	11200 (13.2)	3064 (3.6)	17034 (20.1)	13416 (15.8)	28760 (34.0)	11215 (13.2)
mCCI							
0	1187965	21890 (1.8)	2908 (0.2)	22287 (1.9)	86730 (7.3)	769100 (64.7)	285050 (24.0)
1	229651	7981 (3.5)	1273 (0.6)	9272 (4.0)	28553 (12.4)	128966 (56.2)	53606 (23.3)
2	51295	2956 (5.8)	567 (1.1)	6211 (12.1)	9787 (19.1)	22698 (44.2)	9076 (17.7)
3+	15686	1220 (7.8)	253 (1.6)	3069 (19.6)	3113 (19.8)	6068 (38.7)	1963 (12.5)
Men							
Aged 25-44	815950	5659 (0.7)	461 (0.1)	4205 (0.5)	25050 (3.1)	614615 (75.3)	165960 (20.3)
Aged 45-64	458384	19595 (4.3)	2105 (0.5)	12211 (2.7)	86437 (18.9)	234266 (51.2)	103770 (22.6)
Aged 65-74	96404	9870 (10.2)	1607 (1.7)	9572 (9.9)	26821 (27.8)	31910 (33.1)	16624 (17.2)
Aged 75-84	49438	7551 (15.3)	2298 (4.6)	12238 (24.8)	7174 (14.5)	14630 (29.6)	5547 (11.2)
mCCI							
0	1173065	30524 (2.6)	4269 (0.4)	22906 (2.0)	104942 (8.9)	763831 (65.1)	246593 (21.0)
1	201200	8228 (4.1)	1368 (0.7)	7903 (3.9)	29919 (14.9)	113921 (56.6)	39861 (19.8)
2	34665	2814 (8.1)	549 (1.6)	4758 (13.7)	8088 (23.3)	13994 (40.4)	4462 (12.9)
3+	11246	1109 (9.9)	285 (2.5)	2659 (23.6)	2533 (22.5)	3675 (32.7)	985 (8.8)

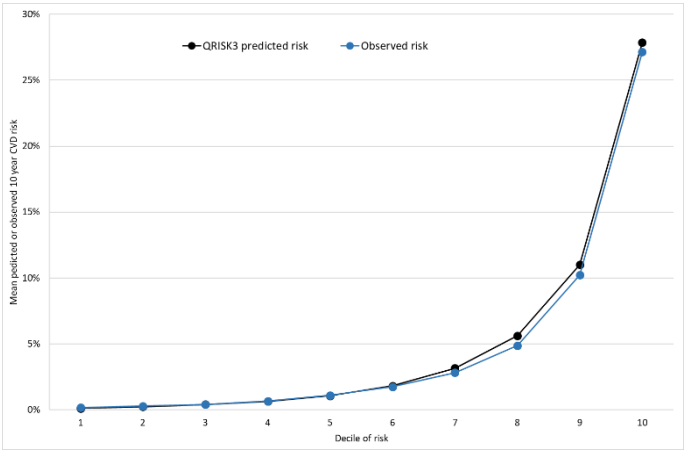
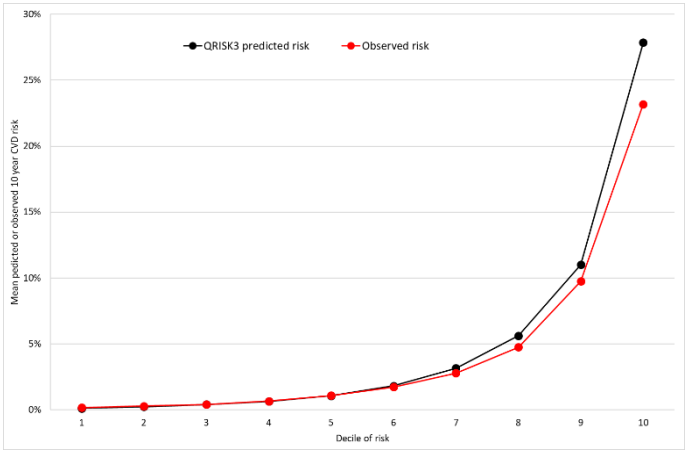
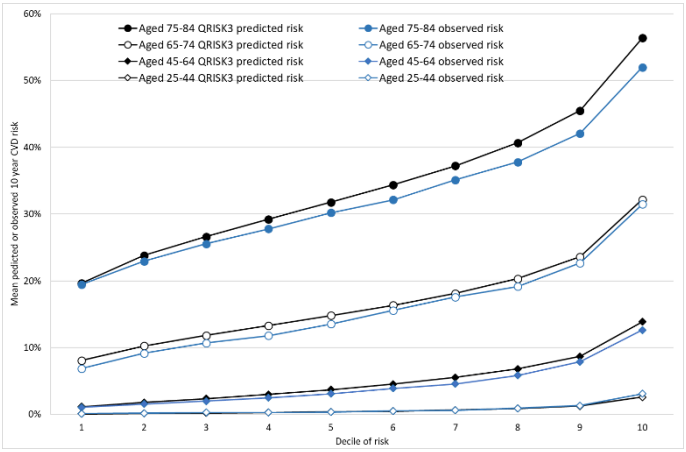
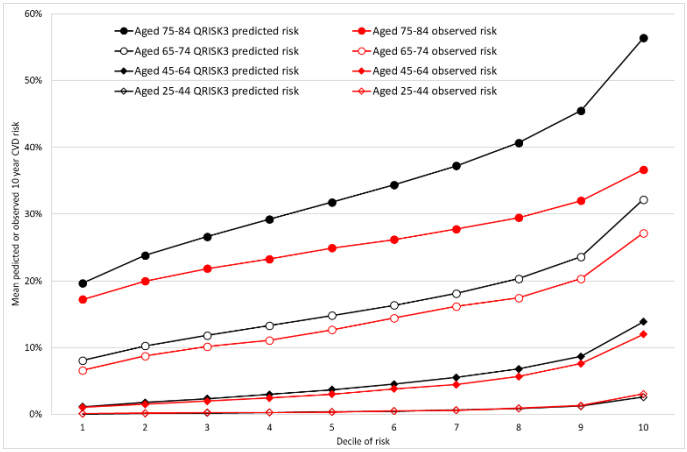
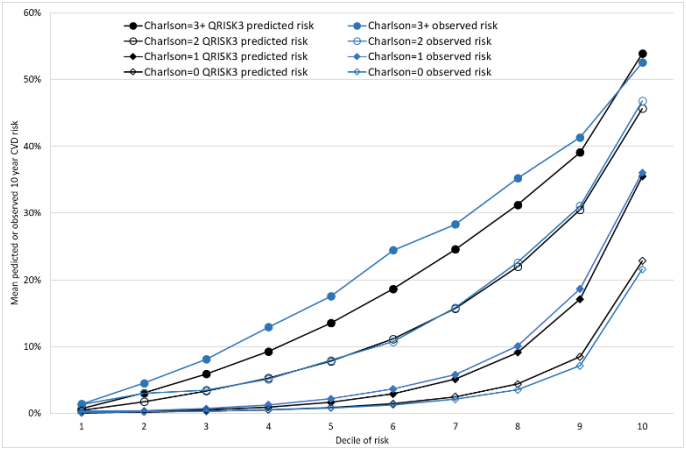
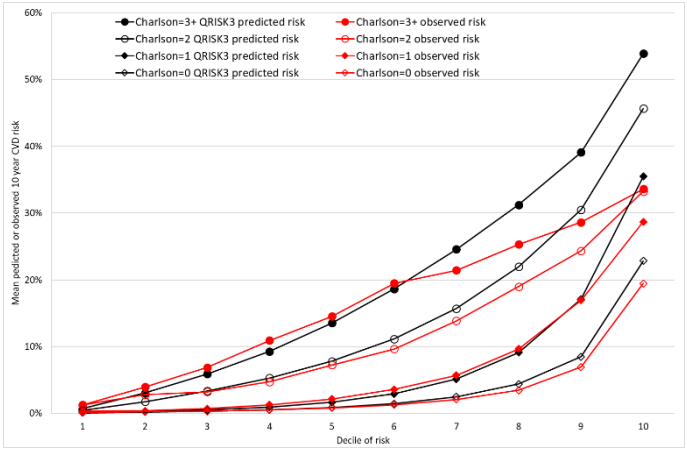
mCCI: modified Charlson Comorbidity Index

Table 3: Discrimination and model fit

	Women			Men		
	Harrell's C	D	R-squared	Harrell's C	D	R-squared
All patients	0.865 (0.861,0.868)	2.43 (2.41,2.45)	58.5(58.1,58.8)	0.834 (0.831,0.837)	2.10 (2.08,2.12)	51.3(50.8,51.7)
Age-group						
25-44	0.758 (0.747,0.69)	1.69 (1.63,1.76)	40.7 (38.8,42.5)	0.757 (0.749,0.764)	1.57 (1.52,1.61)	36.9 (35.6,38.2)
45-64	0.707 (0.702,0.713)	1.25 (1.22,1.28)	27.2 (26.1,28.3)	0.681 (0.677,0.685)	1.04 (1.02,1.07)	20.6 (19.8,21.4)
65-74	0.641 (0.635,0.647)	0.82 (0.77,0.86)	13.7 (12.4,15.1)	0.612 (0.606,0.617)	0.63 (0.59,0.66)	8.6 (7.7,9.5)
75-84	0.611 (0.605,0.616)	0.61 (0.56,0.66)	8.1 (6.9,9.3)	0.585 (0.579,0.591)	0.46 (0.42,0.51)	4.9 (4.1,5.8)
mCCI						
0	0.863 (0.859,0.867)	2.40 (2.38,2.43)	57.9 (57.4,58.4)	0.827 (0.824,0.831)	2.02 (2.00,2.04)	49.4 (48.9,49.8)
1	0.846 (0.840,0.852)	2.20 (2.17,2.24)	53.6 (52.8,54.4)	0.829 (0.823,0.835)	2.00 (1.96,2.03)	48.7 (47.8,49.6)
2	0.789 (0.778,0.799)	1.73 (1.67,1.78)	41.6 (39.9,43.2)	0.728 (0.717,0.739)	1.28 (1.22,1.34)	28.1 (26.2,29.9)
3 or more	0.744 (0.728,0.760)	1.40 (1.32,1.48)	31.8 (29.2,34.4))	0.695 (0.678,0.712)	1.13 (1.04,1.21)	23.2 (20.5,26.0)

mCCI: modified Charlson Comorbidity Index.

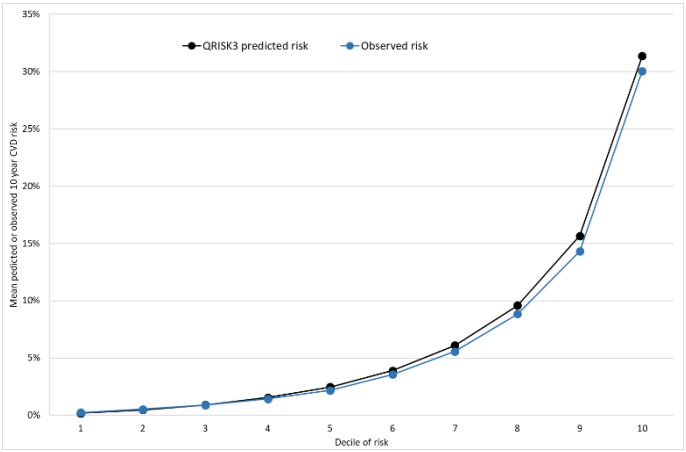
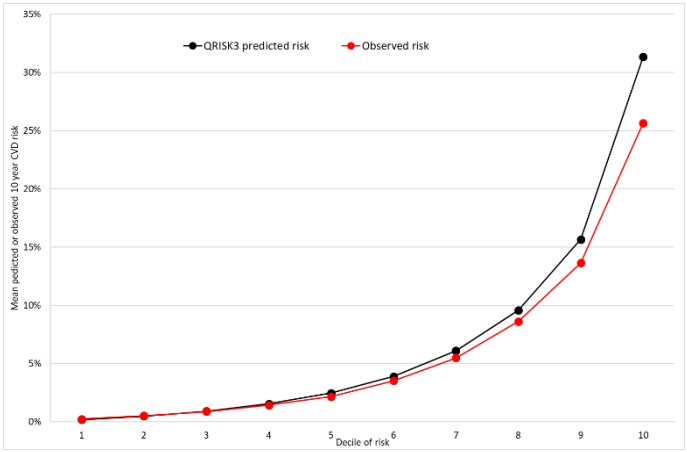
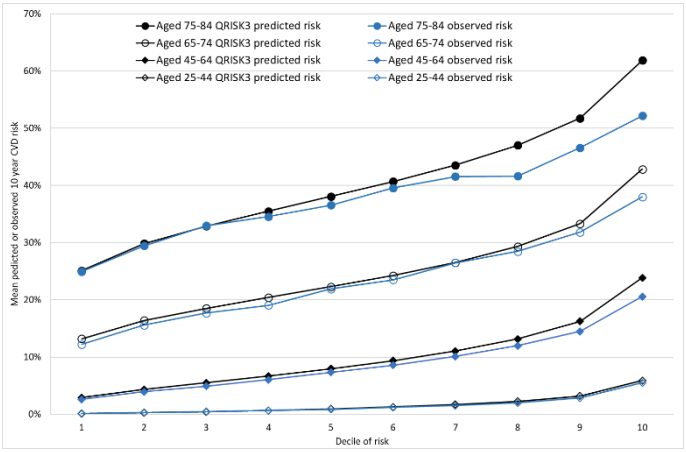
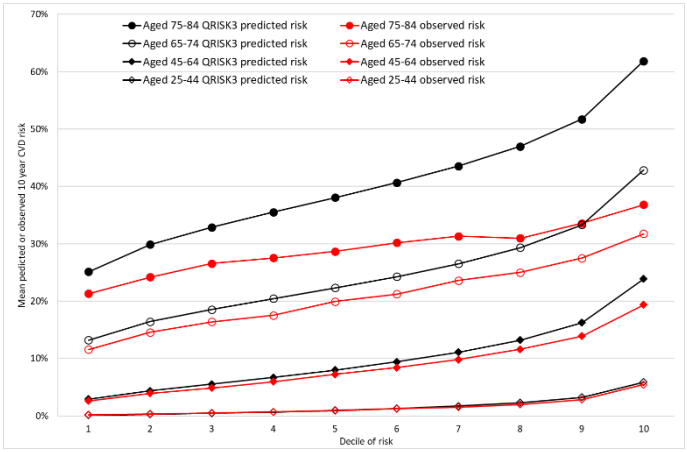
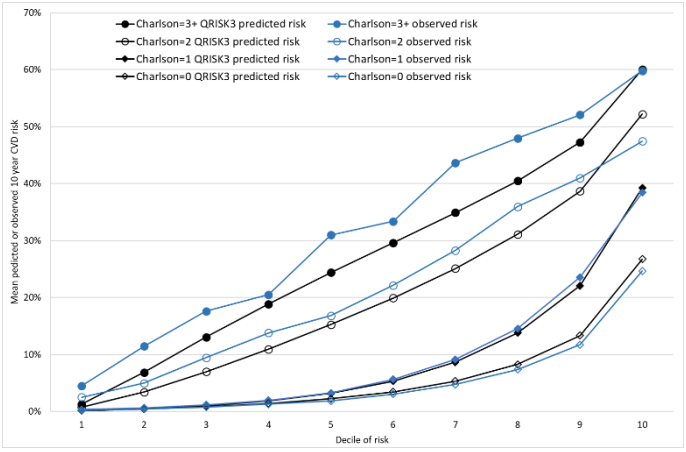
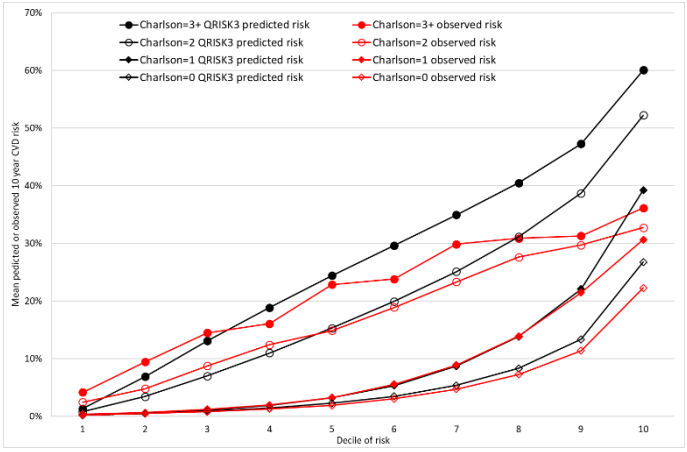
Figure 1: Calibration in women without accounting for competing risks (left hand) and accounting for competing risks (right hand)

Women - not accounting for competing risks*	**Women - accounting for competing risks#**
1a: Overall calibration	1b: Overall calibration
	
1c: Calibration by age group	1d: Calibration by age group
	
1e: Calibration by Charlson Comorbidity Index	1f: Calibration by Charlson Comorbidity Index
	

* Observed risk is based on the Kaplan-Meier estimator, which does not account for competing mortality risk.

Observed risk is based on the Aalen-Johansen estimator, which accounts for competing mortality risk

Figure 2: Calibration in men without accounting for competing risks (left hand) and accounting for competing risks (right hand)

Men - not accounting for competing risks*	**Men - accounting for competing risks#**
2a: Overall calibration	2b: Overall calibration
	
2c: Calibration by age group	2d: Calibration by age group
	
2e: Calibration by Charlson Comorbidity Index	2f: Calibration by Charlson Comorbidity Index
	

* Observed risk is based on the Kaplan-Meier estimator, which does not account for competing mortality risk.

Observed risk is based on the Aalen-Johansen estimator, which accounts for competing mortality risk

References

1. National Institute for Health and Care Excellence. Clinical Guideline 181: Lipid modification: cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease. London: National Institute for Health and Care Excellence; 2014.
2. Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults. Circulation 2014;129:S1-S45.
3. Collins GS, Altman DG. Predicting the 10 year risk of cardiovascular disease in the United Kingdom: independent and external validation of an updated version of QRISK2. BMJ 2012;344:e4181.
4. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ 2017;357.
5. Steyerberg E. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.
6. Collins GS, de Groot JA, Dutton S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. BMC Medical Research Methodology 2014;14:40.
7. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Medical Research Methodology 2013;13:33.
8. van Staa T-P, Gulliford M, Ng ESW, Goldacre B, Smeeth L. Prediction of Cardiovascular Risk Using Framingham, ASSIGN and QRISK2: How Well Do They Predict Individual Rather than Population Risk? PLoS ONE 2014;9:e106455.
9. Wolbers M, Koller MT, Witteman JCM, Steyerberg EW. Prognostic Models With Competing Risks: Methods and Application to Coronary Risk Prediction. Epidemiology 2009;20:555-61.
10. Nguyen Quoc D, Odden Michelle C, Peralta Carmen A, Kim Dae H. Predicting Risk of Atherosclerotic Cardiovascular Disease Using Pooled Cohort Equations in Older Adults With Frailty, Multimorbidity, and Competing Risks. Journal of the American Heart Association 2020;9:e016003.
11. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). International Journal of Epidemiology 2015;44:827-36.
12. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. British Journal of Clinical Pharmacology 2010;69:4-14.
13. Khan N, Perera R, Harper S, Rose P. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. BMC Family Practice 2010;11:1.
14. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 2011;45:67.
15. Rubin D. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley and Sons; 1987.
16. Gerds TA, Kattan MW, Schumacher M, Yu C. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. Statistics in Medicine 2013;32:2173-84.
17. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. Statistics in Medicine 2004;23:723-48.
18. Altman D, Vergouwe Y, Royston P, Moons K. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.
19. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Statistics in Medicine 2007;26:2389-430.

20. Collins G, Reitsma J, Altman D, Moons K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. BMJ 2015;350:g7594.
21. Pylypchuk R, Wells S, Kerr A, et al. Cardiovascular disease risk prediction equations in 400,000 primary care patients in New Zealand: a derivation and validation study. The Lancet 2018;391:1897-907.
22. Koller MT, Leening MJG, Wolbers M, et al. Development and Validation of a Coronary Risk Prediction Model for Older U.S. and European Persons in the Cardiovascular Health Study and the Rotterdam Study. Annals of Internal Medicine 2012;157:389-97.
23. Li Y, Sperrin M, Ashcroft DM, van Staa TP. Consistency of variety of machine learning and statistical models in predicting clinical risks of individual patients: longitudinal cohort study using cardiovascular disease as exemplar. BMJ 2020;371:m3919.
24. Melberg T, Nygård OK, Kuiper KK, Nordrehaug JE. Competing risk analysis of events 10 years after revascularization. Scandinavian cardiovascular journal : SCJ 2010;44:279-88.
25. Ashburner JM, Go AS, Chang Y, et al. Influence of Competing Risks on Estimating the Expected Benefit of Warfarin in Individuals with Atrial Fibrillation Not Currently Taking Anticoagulants: The Anticoagulation and Risk Factors in Atrial Fibrillation Study. J Am Geriatr Soc 2017;65:35-41.
26. Abdel-Qadir H, Fang J, Lee DS, et al. Importance of Considering Competing Risks in Time-to-Event Analyses: Application to Stroke Risk in a Retrospective Cohort Study of Elderly Patients With Atrial Fibrillation. Circ Cardiovasc Qual Outcomes 2018;11:e004580.
27. Read SH, van Diepen M, Colhoun HM, et al. Performance of Cardiovascular Disease Risk Scores in People Diagnosed With Type 2 Diabetes: External Validation Using Data From the National Scottish Diabetes Register. Diabetes Care 2018.
28. Li Y, Sperrin M, Belmonte M, Pate A, Ashcroft DM, van Staa TP. Do population-level risk prediction models that use routinely collected health data reliably predict individual risks? Scientific Reports 2019;9:11222.